Advanced Topics in STR DNA Analysis

Capillary Electrophoresis Instrumentation: Theory and Application

There are 23 pairs of Chromosomes

3 billion base pairs means \( \frac{1}{(\frac{1}{4})^{3,000,000,000}} \)
possible combinations

DNA Analysis by RFLP

The Process of DNA Typing via PCR

Inventor of the PCR

"I think I might have been stupid, in some respects, if it weren't for my psychedelic experiences.

-Kary Mullis, Ph.D., Nobel Prize Laureate, Chemistry, 1993

http://www.cstl.nist.gov/biotech/strbase/training.htm
The Application

- Speed and detection capabilities of DNA analyses have improved since the development of PCR
  - Increase in number of complex assays necessitates automated testing procedures
- Automated systems are needed to increased sample throughput
  - Automated systems must be robust and must demonstrate long term stability

What are the keys to a useful measure of genetic variability, esp. with STRs?

- Reproducible results from day to day
- Resolution of a single base over the range of analysis
- Precision under 0.17 bp for size separation
- Stability over time and insensitivity to matrix effects
- Relative accuracy (not absolute)

Methods of determination of genetic variability

- Probe hybridization
- Charge based mobility and separation – gel and capillary electrophoresis
- Partitioning and ion exchange – HPLC
- Conformation – SSCP, heteroduplex polymorphism
- Size measurement – Mass Spectrometry
  - All of these have been used one time or another for STR/VNTR analysis

How do the various methods add up at present?

- Probe based methods-
  - can be difficult to detect length variations
- HPLC-
  - lacks resolution
- MS –
  - has trouble with sizes above 90bp
- Conformational polymorphisms-
  - will not always vary sufficiently
- Electrophoresis-
  - currently best option- but can have trouble with precision and resolution

The Issues

1. Although the PCR is rapid and efficient, sample loads keep increasing
2. Soon all sexual offenders (and other felons) will be required to submit a sample for testing. Current estimated backlog is 540,000 samples.
3. The number of untested rape kits nationwide is estimated to be 180,000 to 500,000.
4. What technique could be used to automate the analysis of so many samples?
Why Use Capillary Electrophoresis for DNA Analysis?

1. Injection, separation, and detection are automated.
2. Rapid separations are possible
3. Peak information is automatically stored for easy retrieval.

Process Involved in 310/3100 Analysis

- **Injection**
  - Electrokinetic injection process (formamide, water)
  - Importance of sample stacking
- **Separation**
  - Capillary – 50um fused silica, 47 cm (36 cm to detector)
  - POP-4 polymer – Polymethyl acrylamide
  - Buffer - TAPS pH 8.0
  - Denaturants – urea, pyrolidinone
- **Detection**
  - Fluorescent dyes with excitation and emission traits
  - CCD with defined virtual filters produced by assigning certain pixels

Electrophoresis Theory

P = VI = IR
\( v_{ep} = \mu_{ep} V \) Ions move through pickle faster at high voltage
\( \mu_{ep} = \frac{q}{6\pi \eta r} \) Small ions with high charge move fastest

DNA and Electrophoresis

“From a practical point of view it is disappointing that electrophoresis cannot be used to fractionate or analyze DNA on the basis of size” Olivera, Biopolymers 1964, 2, 245

\( \mu_{ep} = \frac{q}{6\pi \eta r} \) small ions with high charge move fastest

As size increases so does charge!
Separation Mechanism

Electrophoretic flow

DNA

Electroosmotic flow

DNA

Issues with CE separations

- Effect – electroosmotic flow

Polyethylene oxide separation of pBR 322 HAE III digest (EOF present)

PDMA (POP) separation of DNA (EOF not present)

In the early 1990s the real question was how to transition from a gel to a capillary

- Cross-linked acrylamide gel filled capillaries were tried first
  - Reusable?
  - Bubble formation
  - Thermal degradation

- Alternative was to not use a gel at all
  - Refillable sieving polymers
  - However, resolution initially was poor with these polymers

So what are sieving buffers?

They are gels - very similar to polyacrylamide

They are not gels - they flow

Actually these are known as entangled linear polymers and there are many common applications

Entangled Polymer Solutions

- Polymers are not cross-linked (above entanglement threshold)
- “Gel” is not attached to the capillary wall
- Pumpable – can be replaced after each run
- Polymer length and concentration determine the separation characteristics
- Examples:
  - 1% HEC (hydroxyethyl cellulose)
  - 4% polyvinyl pyrrolidinone
  - POP-4 and POP-6

Transient Pores Are Formed Above the Entanglement Threshold.

C < C*  C = C*  C > C*
The electric field strength can influence the shape of the DNA molecule.

The Keys
1. Polymer strand interactions create pores
2. Average pore size ~ average DNA volume
3. Viscosity should be minimized
4. Field strength optimized
How to Improve Resolution?
1. Lower Field Strength
2. Increase Polymer Concentration
3. Increase Polymer Length
4. Use a longer capillary

All of these come at a cost of longer separation run times

How Are Separations Performed in a 310?
We use POP-4
(4% poly-dimethylacrylamide, 8M urea, 5% pyrrolidinone)


Running buffer contains
100 mM TAPS and 1 mM EDTA (adjusted to pH 8.0 with NaOH)

Synthesis Procedure for PDMA
(Molecular Wt = 1 Million amu)

- Distill dimethyl acrylamide to remove stabilizers
- Add 16.3 ml of methanol to 46.3 ml dH2O
- Added 6.3 g of dimethyl acrylamide to mixture
- N₂ bubbled through for 1 h (covered flask to prevent excess methanol evaporation)
- Add 0.3 ml of ammonium persulfate stock solution (made by dissolving 0.2 g of APS in 1.8 ml of dH2O) to the methanol/H₂O mixture
- Remove solvents and dry to powder


Synthesis Results
Effect of Concentration and Molecular Weight on resolution

4% PDMA (100K), Taps buffer
7.3% PDMA (1M), Taps buffer

Separation Issues
- Capillary wall coating – dynamic coating with polymer
  - Wall charges are masked by methyl acrylamide
- Electrophoresis buffer –
  - Urea for denaturing and viscosity
  - Buffer for consistent pH
  - Pyrrolidinone for denaturing DNA
  - EDTA for stability and chelating metals
- Polymer solution –
  - Entangled to separate DNA
  - High molecular weight for good resolution
  - Minimum concentration/viscosity for easy refilling (POP4, POP6)
- Run temperature – 60 °C helps reduce secondary structure on DNA and improves precision. (Temperature control affects DNA sizing)

Commercial POP-4
- 4% poly(dimethylacrylamide) (PDMA),
- 100 mM TAPS (pH 8.0), 8 M urea, 5% 2-pyrrolidinone

http://www.cstl.nist.gov/biotech/strbase/training.htm
Injection

How is injection accomplished on a 310

Injection Methods for CE

Stacking Effects

(a) Stacking with Low Ionic Strength

(b) Regular EK Injection

Ion Mobility Effects

The Injection of DNA by voltage is described by

$$[\text{DNA}_{\text{inj}}] = E(\pi r^2)[\text{DNA}_{\text{sam}}](\mu_{\text{ep}} + \mu_{\text{eof}})$$

However this equation assumes no interfering ions are present.

Cl⁻ ions and other interferents will compete with DNA

$$\{\text{DNA}_{\text{inj}}\} = [\text{DNA}_{\text{sam}}][\text{other ions}]$$

Ions such as Cl⁻ have a higher charge/mass ratio and $\mu_{\text{ep}}$ is greater

http://www.cstl.nist.gov/biotech/strbase/training.htm
Two Major Effects of Sample Stacking
1. Sample is preconcentrated. Effect is inversely proportional to ionic strength
2. Sample is focused. Ions stop moving in low electric field
3. Mobility of sample = $\mu_e = \frac{\text{velocity}}{\text{electric field}}$

Comments on Sample Preparation
- Use high quality formamide (<100 $\mu$S/cm)!
  - ABI sells Hi-Di formamide
  - regular formamide can be made more pure with ion exchange resin (or less! You better measure it, aliquot it out, and freeze it!)
- Deionized water vs. formamide
  - water works fine but samples are not stable as long as with formamide, water also evaporates over time...
- Denaturation with heating and snap cooling
  - use a thermal cycler for heating and cold aluminum block for snap cooling
  - heat/cool denaturation step is necessary only if water is substituted for formamide...

Effect of Formamide on Peak Resolution and Sensitivity (Rox Internal Standard)

<table>
<thead>
<tr>
<th>Solvent</th>
<th>Resolution</th>
<th>Peak Height</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>1.19 +/- 0.01</td>
<td>2700 +/- 300</td>
</tr>
<tr>
<td>Formamide (27 $\mu$S)</td>
<td>1.15 +/- 0.05</td>
<td>2960 +/- 30</td>
</tr>
<tr>
<td>Formamide (360 $\mu$S)</td>
<td>1.20 +/- 0.08</td>
<td>879 +/- 4</td>
</tr>
<tr>
<td>Formamide 1000 $\mu$S)</td>
<td>1.20 +/- 0.06</td>
<td>290 +/- 14</td>
</tr>
</tbody>
</table>

Typical Sample Preparation for ssDNA
1. Perform PCR with dye-labeled primers
2. Dilute 1 $\mu$L PCR product with 24 $\mu$L deionized formamide; add 1 $\mu$L ROX-labeled internal sizing standard
3. Denature 2 minutes at 95°C with therocycler
4. Cool to 4°C in therocycler or ice bath
5. Sample will remain denatured for at least 3 days

Injection Study
Evaluate the effects of sample injection on electrophoretic separations by CE.
- different solvents (water and formamide of varying purity);
- different concentration of the sample;
- addition of salts;
- sample stacking

Electrokinetic injection has some unusual properties!

Effect of increasing relative sample volume
Sensitivity improves only in poor quality formamide!

As [DNA] increases so does ionic strength, making injection almost independent of $\mu$L DNA injected

http://www.cstl.nist.gov/biotech/strbase/training.htm
Effect of Increasing salt concentration
(Reannealed ds DNA moves faster)

Detection

Detection Issues

- Fluorescent dyes
  - spectral emission overlap
  - relative levels on primers used to label PCR products
  - dye "blobs" (free dye)
- Virtual filters (determine which pixels are used
  - hardware (CCD camera)
  - software (color matrix)

http://www.cstl.nist.gov/biotech/strbase/training.htm
Laser Used in ABI 310

- Argon Ion Laser
- 488 nm and 514.5 nm for excitation of dyes
- 10 mW power
- Lifetime ~5,000 hours (1 year of full-time use)
- Cost to replace ~$5,500
- Leads to highest degree of variability between instruments and is most replaced part
- Color separation matrix is specific to laser used on the instrument
- Laser is on unless instrument is off!

Methods for Fluorescently Labeling DNA

- Intercalating Dyes (post-PCR)
- Dye-labeled nucleotide insertion during PCR
- Dye-labeled primer insertion during PCR

Fluorescent Emission Spectra for ABI Dyes

- 5-FAM
- JOE
- NED
- ROX

Amine Reactive Dyes used in Labeling DNA

The succinimidyl ester reacts rapidly with amine linkers on DNA bases

Virtual Filters Used in ABI 310

- Visible spectrum range seen in CCD camera
- Commonly used fluorescent dyes
- Arrows indicate the dye emission spectrum maximum

Please Note!

- There are no filters in a 310
- It's just the choice of pixels in the CCD detector
- All the light from the grating is collected
- You just turn some pixels on and some off
Why Make a Matrix?

The matrix is the solution to a problem:
What’s the contribution at any given wavelength (filter set) from each dye?

There are 4 dyes

- Remember algebra from high school?
- To solve a problem with 4 unknowns, you need 4 equations

Matrix Standards (Raw Data)

<table>
<thead>
<tr>
<th>Filter Set C</th>
<th>Set E</th>
</tr>
</thead>
<tbody>
<tr>
<td>6FAM (5FAM)</td>
<td></td>
</tr>
<tr>
<td>TET (JOE)</td>
<td></td>
</tr>
<tr>
<td>HEX (NED)</td>
<td></td>
</tr>
<tr>
<td>ROX (ROX)</td>
<td></td>
</tr>
</tbody>
</table>

For Example

\[
\begin{align*}
I_{540} &= bx_b + gy_b + yz_b + rw_b \\
I_{560} &= bx_g + gy_g + yz_g + rw_g \\
I_{580} &= bx_y + gy_y + yz_y + rw_y \\
I_{610} &= bx_r + gy_r + yz_r + yw_r
\end{align*}
\]

Where

- \(b\) is the %blue labeled DNA
- \(g\) is the %green labeled DNA, etc.
- \(x, y, z, w\) are the numbers in the matrix (sensitivity to each color)

If you solve \(xyzw\) for each dye individually
Then you can determine dye contribution for any mixture

DNA Size Estimation with ABI 310

Once you can label DNA you must next determine its size

1. Each sample is run with a ROX internal standard
2. An external standard is run with ROX as well
3. The unknown allele sequence is determined by comparison to the known ladder allele
4. Assumptions?

http://www.cstl.nist.gov/biotech/strbase/training.htm
Estimating size

The red internal lane standard establishes the allele size. This size is compared to an allelic ladder run earlier.

Assumptions with ABI 310 Method affecting precision

1. DNA is a sphere. (it is not)
2. The conditions for unknown run are the same as the ladder run. (they are not)
3. The ROX dye migrates relatively the same as the FAM dye. (It does not)
4. A calibration for one ladder is good for an entire run (sometimes)
5. Temperature is constant (to what degree?)

There is a size range for calling an allele

These bins define the precision of the system

They are commonly defined as +/- 0.5 bp

However they can also be defined by the precision of an allele call (+/- 3 standard deviations)

What affects precision?

Lots of things:

- Temperature
- Sequence of Rox standard vs sample
- Sequence of allele vs ladder
- Conformation of DNA
- Polymer matrix
- Capillary condition
- Buffer concentration
- pH

- Showing that you can control these factors is the goal of laboratory validation
- That will be discussed in future sections.

Conclusions

DNA typing by capillary electrophoresis involves:

1) The use of entangled polymer buffers
2) Injection by sample stacking
3) Multichannel laser induced fluorescence
4) Internal and external calibration

McCord Research Group
Ohio

McCord Research Group
Miami
Acknowledgements

Funding: National Institute of Justice

National Science Foundation

People:

- Ohio University
- Tim Nock
- Jim Dove
- Sharon Williams
- Dr. Federica Crivellente
- Kylie Muncy
- Dr. Brittany Hartzel
- Dr. Yin Shen
- Dr. Nancy Tatarak
- Dr. Denise Chung

- Elsewheres
- Dr. Alice Isenberg FBI
- Dr. John Butler NIST
- Dr. Ralph Allen UVA
- Dr. David Mao J&W Scientific
- Dr. John Petersen UT-Galveston
- Dr. Lee Meadows Jantz

- Florida International University
- Kerry Opel
- Rayna Hebard
- Sarah Hughes
- Dr. Dee Mills

Points of view expressed in this presentation are those of the authors and do not necessarily represent the official view of the US department of Justice.